3段階相対基準型ブロック方式の予選審査設計問題

——マネジメントサイエンスの適用(1)——

岡　田　　勇

1　はじめに

本稿で取り扱う予選審査とは，事前には評価対象数が不明であり，一意に評価が定まらないため，複数の審査員による評価の統合が必要なコンテストの一部である。こういったコンテストには，公募作品のコンテスト，入社試験，音楽のコンクールなどが該当するため，適用可能性が広く，本研究の知見も広く一般化できる。コンテストは参加するチーム数が未定であるうえ，定性的評価を行う関係から複数の審査員を確保しなければならず，また審査時間に制約があるなど，その制度設計は複雑になる場合が多い。そのため，コンテストの設計問題は，オペレーションズリサーチ，あるいは，マネジメントサイエンスの適用例として位置づけることができる。

本稿では，以下のようなコンテストを想定している。参加するチーム数が少ない場合は，1段階審査のみとし，ある程度の審査員数による評価を行うだけでよいが，チーム数がある程度を超えると2段階審査を行う必要がある[1]。この場合，1段階目の審査を予選，2段階目の審査を本選と呼ぶ。コンテストのうち，上位入賞者を表彰する場合は，本選はより精度の高い審査を行う必要がある。一方，予選は，その結果を本選に反映させない場合は，本選出場に対する合否判定が審査の主要な目的となるため，本選に比べて，より容易に行うことができる。さらに，コストを含め審査員の確保は厳しい制約になりやすく，審査の妥当・公平性を確保できる最低限の審査員数で制度設計する必要がある。また，審査にかかる時間を考慮すると，一度に行えるチーム数にも上限がある。

予選審査をどのように行うかは制度設計の重要なテーマであるが，本稿では，ブロック方式について検討する。すなわち，全ての参加するチーム数をいくつかのブロックに分け，各ブロックで独立に予選を行い，それぞれブロックから本選出場チームが選ばれるという方式である。1ブロック当たりのチーム数を2，各ブロックから選ばれるチーム数を1としたとき，トーナメント方式となる。この方式は審査の負担が軽減されるメリットがあり，審査員制約が強い事例ではよく採用されている。

ブロック方式を採用した場合，強豪チームが同一ブロックに配当されてしまう不公平をいかに

1　もちろん，それでもチーム数が多い場合は，3段階以上の審査が必要になるが，本稿では扱わない。

回避するかが問題となる。ある程度実力を事前予測でき，ブロックに分ける段階で公平な操作ができる場合を除くと，この欠点を回避することは困難である。また，チームの評価を100点満点でいくらであるといった絶対的な評価が可能であれば，どのようなブロック分けであっても，事実上ブロックをまたがった全体審査を行っていることになるので，不公平感は生じない。これを絶対基準型ブロック方式と呼ぶが，本研究ではこのような評価が不可能であるケースを考える。

絶対基準型でない評価ルールである相対基準型では，審査員はあくまでもブロック内での相対的な順位に基づく評価にとどまる。このとき，予選審査は通過判定であるので，審査員に詳細な順位尺度を要求するのは不自然である。そのため，審査員には個々のチームの合否判定かそれに補欠合格を認めるような，高々３段階（合格・補欠合格・不合格）程度の評価を行うのが妥当であろう[2]。

このような判定を審査員に課す時，予選通過数を固定するような制度の設計は困難である。審査員から受けた判定の総合点が同一のチームが複数存在するときに，もう一度審査を行う必要があるからである。これを完全に回避するには理論的には全てのチームを順位づける必要があるが，これは前段の通り取り扱わないこととする。すなわち，３段階相対基準型ブロック方式の場合，予選通過数を事前に決定できない[3]。しかし，本選出場チーム数にも一定の制約があるのが通常であることを考えると，なるべく予選通過数を予測可能にしておいた方がよい。そのため本稿では，個々の審査員が判定する合格・補欠合格・不合格のチーム数をあらかじめ固定する方法について検討する。

審査員の判定チーム数を固定する方法の特徴はどうであろうか。このような固定ルールを採用すると，現実のチーム評価に困る場合が生じる。例えば実力が伯仲するチームが３つあったときに，合格チームを２つと決められていることは審査員の評価の困難性を増大させる。この究極が全てのチームの順位付けである。しかし，全てのチームの順位付けほど評価は困難ではない。つまり，この方法は，審査員の評価の困難さと本選出場チームの制御可能性とのトレードオフとしてとらえるべきであろう。

以上の考察から，３段階相対基準型ブロック方式の位置づけが明確になった。本稿では，この方式で予選審査を行うコンテストの設計問題を取り上げる。２節では，検討する方式をより具体的に定式化する。３節では問題の本質的な数理モデルを抽象化し解析を行う。４節では，解析結果に基づき予選審査の制度設計を行う。５節では，シミュレーションを用いて制度の評価を行う。そして６節で本論をまとめる。

2　より詳細な順位づけを行う事例も存在するが，本稿では取り扱わない。

3　チーム数を固定しないことは別のメリットも存在するかもしれない。一般的に実力が伯仲するチームのときに，ともに本選出場を果たすといったことが可能となるため，ブロック編成の不公平感をある程度減少させるかもしれない。

2　問題の定義

ここでは，以下のような制約と定式化を仮定する[4]。参加するチーム数Nは事前には未定であるが，経験的に$5 \leq N \leq 55$とする。審査時間の制約から一回に行うチーム数に上限があるため，本選出場チーム数N_2は$8 \leq N_2 \leq 12$が望ましいものとする。また予選で行うブロック当たりのチーム数N_1は$3 \leq N_1 \leq 8$とする。ここでN_1は制約条件とするが，N_2は相対基準型ブロック方式のため，制約条件ではなく目標値となる。

審査員の数はなるべく少なくしたい。本選は順位を付ける必要があるとすると，審査員の全ての対象への順位付けを統合する評価ルールが必要になる。そのときになるべく多くの審査員がいる方が，評価が割れるようなときの評価の信頼性を担保できる。ここでは本選の審査員の数J_2を$J_2 = 7$とする。予選は合否判定のみを行えばよいので，少ない人数で対応できる。ここでは，予選の審査員の数J_1を$J_1 = 3$とする。

予選通過基準をどう設定すべきか。そのため，予選における審査員の個々のチームの判定を合否判定，それを統合して予選通過か否かを判定する基準を予選通過基準と呼ぶ。我々は，過半数の審査員が合格と判定したチームを予選通過とするのが妥当であると考えるので，それを以下のように定式化する。

予選通過基準 $C : P_i = 1$，iff $\sum_j J_{ij} > J_1/2$

ただし，P_iとは予選通過判定を表し，チームiが予選通過をする場合は1，そうでない場合は0である。J_{ij}とは審査員の評価を表し，チームiに対する審査員jの判定が合格の時は1，そうでない時は0とする。この基準は過半数ルールに基づく基準であるので，妥当性が高い。また後で補欠合格という概念を導入するが，これを$J_{ij} = 1/2$と設定することで，補欠合格を含んだときの予選通過基準も上記のCをそのまま適用できるメリットがある。この基準Cを適用するための最少人数として$J_1 = 3$を採用している。予選通過するには2人以上の審査員から合格判定が出されなければならない。また補欠合格を採用した場合は，上記に加えて1人の審査員から合格判定が出て，残りの2人の審査員から補欠合格判定が出る必要がある。

次節では，この予選通過基準C に基づくブロック編成の方針を抽象的なモデルの解析から考えたい。

3　問題の抽象化と解析

予選通過基準を具体的に定義するために，問題の本質を抽象化した記述とその解析について議論する。審査員の評価が合格・不合格の2段階であることを抽象化した「等価な選択問題」と，補欠合格を含む3段階であることを抽象化した「補欠を含む選択問題」についてそれぞれ定式化

4　ここでの諸条件は筆者の所属する学部において行われる，学生を対象とする研究論文発表大会を想定している。

を試み，解析を行う。

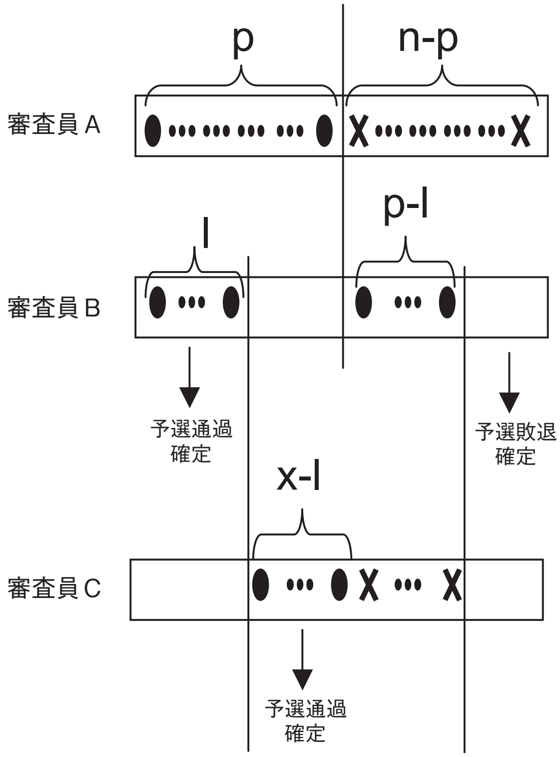


図1：等価な選択問題の期待値に関する証明の概念図

3.1　等価な選択問題

識別可能な n 個の玉から p 個の玉を独立に選ぶ審査員が3人いて，2人以上の審査員が合格と判定した玉を予選通過した玉とする。このとき予選通過した玉が x 個である確率を $P^x(n, p)$ とする。またある玉が予選通過する確率を $P(n, p)$ とする。当然 $\sum_{x=0}^{n} xP^x(n, p) = nP^x(n, p)$ である。ここで，全ての審査員はすべての玉の合格判定を同じ確率で行うものとするとき，以下の定理が成り立つ。

定理1（等価な選択問題の期待値）

$$P(n, p) = \frac{p^2(3n-2p)}{n^3}$$

$$P^x(n, p) = \frac{1}{C(n, p)^2} \sum_{l=max(0,2p-n,3p-n-x,x-p)}^{min(x,p,2p-x)} C(p, l)C(n-p, p-l)$$

$$C(2p-2l, x-l)C(n-2p+2l, p-x+l)$$

証明は次の通り。まず，$P(n, p)$ について。3人から合格と判定される確率は p^3/n^3，2人から合格と判定され1人から不合格とされる確率は $C(3, 2)p^2(n-p)/n^3$ である。

次に，$P^x(n, p)$ について。審査員Aが合格と判定した玉の群 G の中から審査員Bが合格と判定する玉の数を l とする。l は x 以下であるべきであり，G の中から選ばれるので0以上 $|G|=p$ 以下である。また，審査員Bが合格と判定する残りの玉は $\bar{G}$ から $p-l$ 個選ばれることになる。予選通過する玉の数は x 個なので，審査員Cは，自身の選択によって新たに $x-l$ 個の玉が予選通過するように合格判定する必要がある。AとBのどちらか一方からのみ合格とされた玉の群 G' は $|G|=2p-2l$ であるから，審査員Cは G' から $x-l$ 個合格させる必要がある。当然審査員Cは G' からも $\bar{G}'$ からも所定の数だけ玉を合格させなければならない。以上のことから l に関する制約条件として，

$$l \leq x,\ 0 \leq l \leq p,\ p-l \leq n-p,\ x-l \leq 2p-2l,\ 0 \leq p-x+l \leq n-2p+2l$$

となるので，これを満たす l におけるコンビネーションの総和となる。なお概念図を図1に掲げる。

3.2 補欠を含む選択問題

識別可能な n 個の玉から p 個の玉を合格とし，q 個の玉を補欠合格と判定する独立した審査員が3人いる。2人以上の審査員から合格判定を受けるか，1人の審査員から合格判定され，2人の審査員から補欠合格と判定された玉を予選通過した玉と呼ぶ。このとき，予選通過した玉の数が x 個である確率を $P^x(n, p, q)$ とする。またある玉が予選通過する確率を $P(n, p, q)$ とする。当然 $\sum_{x=0}^{n} xP^x(n, p, q) = nP(n, p, q)$ である。また当然，$q=0$ とすると前節の問題を含む。ここで，全ての審査員はすべての玉を合格，もしくは，補欠合格として選ぶ確率は同じものとするとき，前節と同様，以下の定理が成り立つ。

定理2（補欠を含む選択問題の期待値）

$$P(n, p, q) = P(n, p) + \frac{3pq^2}{\mathrm{n}^3}$$

$$\begin{aligned}P^x(n, p, q) = \frac{1}{(C(n, p)C(n-p, q))^2} \sum_{l=\underline{l}}^{\bar{l}} \sum_{l'=\underline{l'}}^{\bar{l'}} \sum_{m=\underline{m}}^{\bar{m}} \sum_{m'=\underline{m'}}^{\overline{m'}} \sum_{k=\underline{k}}^{\bar{k}} \sum_{k'=\underline{k'}}^{\bar{k'}} \\ C(p, l)C(q, l')C(n-p-q, p-l-l')C(p-l, m)C(q-l', m') \\ C(n-2p-q+l+l', q-m-m')C(2p-2l-l'-m+m', k) \\ C(m+l', k')C(n-2p+2l-m', p-k-k') \\ C(m+l'-k', x-l-k-k')C(n-p-m-l'+k', q-x+l+k+k')\end{aligned}$$

ここで，

$\underline{l} = 0,$	$\bar{l} = min(x, p)$
$\underline{l'} = max(0, 2p+q-n-l),$	$\bar{l'} = min(q, p-l)$
$\underline{m} = 0,$	$\bar{m} = p-l$
$\underline{m'} = max(0, 2p+2q-n-m-l-l'),$	$\overline{m'} = min(q-l', q-m)$

$\underline{k}=max(0, x-l-m-l')$,　　　　　　　$\overline{k}=min(x-l, 2p-2l-l'-m+m',$
$n-p-q+x-m-l-l')$

$\underline{k}'=max(0, 3p-n-k-2l+m', x-q-k-l)$,　　$\overline{k}'=min(x-l-k, m+l', p-k)$

である。

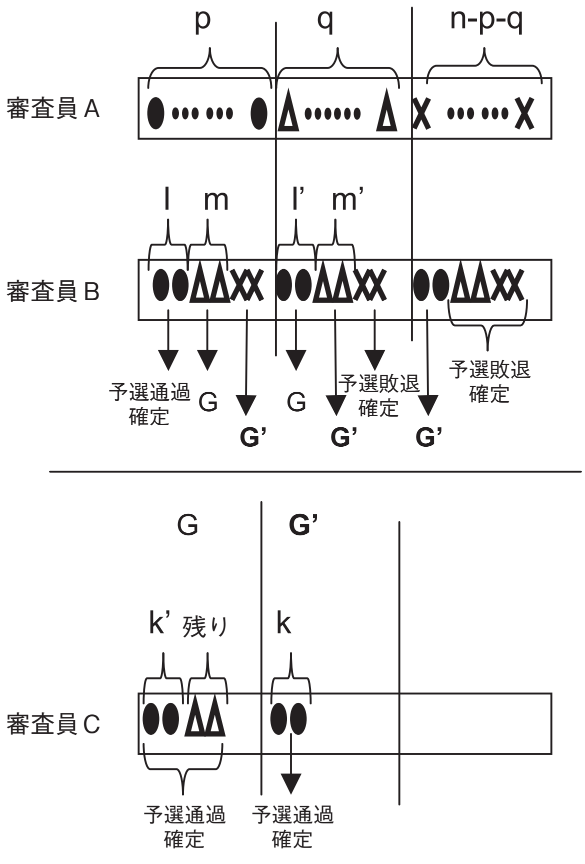


図2：補欠を含む選択問題の期待値に関する証明の概念図

証明を行う。まず $P(n, p, q)$ は3人から合格判定を受ける確率 p^3/n^3，2人から合格と判定され1人から補欠合格とされる確率 $C(3, 2)p^2q/n^3$，2人から合格判定され1人から不合格とされる確率 $C(3, 2)p^2(n-p-q)/n^3$ と，1人から合格判定され2人から補欠合格とされる確率 $C(3, 2)pq^2/n^3$ との和である。

次に $P^x(n, p, q)$ について証明する。審査員Aから合格判定された玉の群を G_A，補欠合格とされた玉の群を G'_A とする。予選通過する玉の数が x なので，x 以下の l 個が審査員Bが G_A の中から合格判定される（H0）。審査員Bが合格判定する残りの玉のうち l' 個が G'_A にある（H1）。このとき審査員Aが不合格と判定した玉の群 $\overline{G_A \cup G'_A}$ から残りの玉が合格判定されることになる（H2）。同様に補欠の玉についても G_A から m 個，G'_A から m' 個選ばれ（H3），

$\overline{G_A \cup G'_A}$ から（Ｈ２）で合格判定された玉を除いた玉の中から $q-m-m'$ 個が選ばれる（Ｈ４）。審査員ＡとＢがともに合格判定した玉の数は l であり，一方が合格でもう一方が補欠合格として判定された玉の群 G_B は $|G_B|=m+l'$ であり，一方が合格でもう一方が不合格判定した玉と，両者とも補欠合格と判定した玉の群 G'_B は $|G'_B|=m'+p-l-m+p-l-l'=2p-2l-l'-m+m'$ である。予選通過した玉の数が x であるから，審査員Ｃは G'_B から k 個合格判定し（Ｈ５），G_B から k' 個合格判定する（Ｈ６）とすると，$x-l-k-k'$ 個を G_B から（Ｈ６）で合格判定された玉を除いた玉の中から補欠合格として選ぶ（Ｈ７）ことになる。審査員Ｃが合格判定とする玉は $\overline{G_B \cup G'_B}$ から選ばれる必要があり（Ｈ８），補欠の玉は G_B 以外の残っている玉の中から選ばれることになる（Ｈ９）。以上のことからそれぞれの性質に対応する制約条件として，

Ｈ０：$l \leq x,\ 0 \leq l \leq p$

Ｈ１：$0 \leq l' \leq q,\ 0 \leq l+l' \leq p$

Ｈ２：$p-l-l' \leq n-p-q$

Ｈ３：$0 \leq m \leq p-l,\ 0 \leq m' \leq q-l',\ 0 \leq m+m' \leq q$

Ｈ４：$q-m-m' \leq n-p-q-(p-l-l')= n-2p-q+l+l'$

Ｈ５：$k \leq x-l,\ 0 \leq k \leq 2p-2l-l'-m+m'$

Ｈ６：$k' \leq x-l-k,\ 0 \leq k' \leq m+l',\ 0 \leq k+k' \leq p$

Ｈ７：$x-l-k-k' \leq m+l'-k',\ 0 \leq x-l-k-k' \leq q$

Ｈ８：$p-k-k' \leq n-(m+l')-(2p-2l-l'-m+m')= n-2p+2l-m'$

Ｈ９：$0 \leq q-x+l+k+k' \leq n-p-(m+l'-k')= n-p-m-l'+k'$

となるので，これを満たす $l,\ l',\ m,\ m',\ k'\ k'$におけるコンビネーションの総和となる。なお概念図を図２に掲げる。

4　予選審査の設計

以上の考察から，予選審査の設計を行う。このため，設計方針を以下のように定める。

設計方針：コンテストに参加するチーム数が確定した段階で，本選出場チーム数が目標値に近づくように，チームをブロックに分ける。このとき各ブロックの本選出場の可能性がなるべく等しくなるように予選審査ルールを設計する。複数の設計が可能であるときは，全ての審査員の合計が少なくなるようにする。

ここで，本選出場の可能性とは，ブロックに所属するチームの実力が同等であるときに，自分のチームが本選出場する確率である。後にシミュレーションにおいて実力に差があるケースを扱うが，本節では実力は同等であるものとして議論する。

設計方針を満足するために，全ての予選審査のルールにおける本選出場の可能性を計算する。表１は $P(n,\ p,\ q)$ を降順に並べた一覧表である。ここで前述の仮定から，$3 \leq n \leq 8$ である。また $0 < p < n,\ 0 \leq q < n-p$ とした。

チーム数によっては複数のルールを組み合わせて適用する必要があるが，設計方針から本選出場の確率が一定以下であるべきである。表１から経験的にそれを0.01以下と設定した。

表１：$P(n, p, q)$ の一覧

n	p	q	P(n, p, q)	n	p	q	P(n, p, q)	n	p	q	P(n, p, q)
8	7	0	0.95703	8	4	1	0.52344	6	2	0	0.25926
7	6	0	0.94461	6	2	3	0.50926	8	1	6	0.25391
6	5	0	0.92593	4	2	0	0.5	7	2	1	0.21574
5	4	0	0.896	6	3	0	0.5	4	1	1	0.20312
8	6	1	0.87891	8	4	0	0.5	8	2	2	0.20312
7	5	1	0.84548	7	3	2	0.49854	5	1	2	0.2
4	3	0	0.84375	7	2	4	0.47813	6	1	3	0.19907
8	6	0	0.84375	8	3	3	0.47461	7	2	0	0.19825
7	5	0	0.80175	8	2	5	0.44922	7	1	4	0.19534
8	5	2	0.80078	7	3	1	0.41983	8	1	5	0.18945
6	4	1	0.7963	5	2	1	0.4	8	2	1	0.16797
7	4	2	0.74636	7	3	0	0.39359	4	1	0	0.15625
3	2	0	0.74074	8	3	2	0.38672	8	2	0	0.15625
6	4	0	0.74074	3	1	1	0.37037	8	1	4	0.13672
5	3	1	0.72	6	2	2	0.37037	7	1	3	0.13411
8	5	1	0.71289	7	2	3	0.35569	6	1	2	0.12963
8	4	3	0.71094	5	2	0	0.352	5	1	1	0.128
8	5	0	0.68359	4	1	2	0.34375	5	1	0	0.104
6	3	2	0.66667	8	2	4	0.34375	8	1	3	0.0957
5	3	0	0.648	8	3	1	0.33398	7	1	2	0.09038
7	4	1	0.6414	5	1	3	0.32	6	1	1	0.08796
7	3	3	0.62974	8	3	0	0.31641	6	1	0	0.07407
7	4	0	0.60641	6	1	4	0.2963	8	1	2	0.06641
8	3	4	0.59766	6	2	1	0.28704	7	1	1	0.06414
4	2	1	0.59375	7	1	5	0.27405	7	1	0	0.05539
8	4	2	0.59375	7	2	2	0.26822	8	1	1	0.04883
5	2	2	0.544	8	2	3	0.26172	8	1	0	0.04297
6	3	1	0.54167	3	1	0	0.25926				

これを用いてチーム数ごとのルールを設計する。本選出場数 N_2 の目標から $N<12$ の時は予選を行わない。$12\leq N\leq 55$ のときに予選を行うものとするが，その時のブロック編成と評価について，前述の設計方針を満足するルールを考える。こうしたときに複数の制度設計が可能である場合も多いが，本論では表２のような制度を考える。表２では，まず表１のうち，$9/N\leq P(n, p, q)\leq 11/N$ を満たし[5]，互いの値の差が0.01 を超えない $P(n, p, q)$ の組を１ないし２つ取り出して列挙した。そのルールの集合 V に対し，

$$\exists\, s, s'\in \aleph : \forall v, v'\in V : sv(n)+s'v'(n)=n,\quad s+s'<8$$

を満たすようなルールの組を選ぶ[6]。最後にその中から，延べ審査員数が最小の組を選ぶとい

5　$43\leq N\leq 47$ の時に限り解が存在しないので，この条件を緩め，$8/N\leq P\leq 11/N$ とした。

うような方針で作成した。表２では，$v00$から $v24$までの25種類のルールが使用され，それぞれ予選出場数 N ごとの採用ルールの組と延べ審査員数[7] J を示している。

表２：予選のためのブロック編成と審査ルール

No.	n	p	q	No.	n	p	q	No.	n	p	q
v00	4	3	0	v09	6	3	0	v18	8	3	0
v01	7	4	2	v10	7	3	2	v19	7	2	2
v02	3	2	0	v11	7	2	4	v20	8	2	3
v03	5	3	1	v12	8	3	3	v21	6	2	0
v04	8	5	0	v13	5	2	1	v22	7	2	1
v05	5	3	0	v14	7	3	0	v23	7	1	4
v06	7	4	1	v15	7	2	3	v24	8	1	5
v07	6	3	1	v16	5	2	0				
v08	4	2	0	v17	5	1	3				

N	J	Rules	N	J	Rules	N	J	Rules
12	16	v00,v00,v00	27	22	v15,v16,v16,v16,v16	42	25	v22,v22,v22,v22,v22,v22
13	16	v01,v02,v02	28	19	v15,v15,v15,v15	43	25	v23,v23,v23,v23,v23,v24
14	13	v01,v01	29	22	v15,v15,v16,v16,v16	44	25	v23,v23,v23,v23,v24,v24
15	16	v03,v03,v03	30	25	v16,v16,v16,v16,v16,v16	45	25	v23,v23,v23,v24,v24,v24
16	13	v04,v04	31	22	v17,v17,v17,v18,v18	46	25	v23,v23,v24,v24,v24,v24
17	16	v05,v05,v06	32	19	v18,v18,v18,v18	47	25	v23,v24,v24,v24,v24,v24
18	16	v07,v07,v07	33	19	v17,v17,v17,v17,v17,v18	48	25	v24,v24,v24,v24,v24,v24
19	16	v09,v09,v10	34	22	v17,v17,v18,v18,v18	49	28	v23,v23,v23,v23,v23,v23,v23
20	22	v08,v08,v08,v08,v08	35	22	v19,v19,v19,v19,v19	50	28	v23,v23,v23,v23,v23,v23,v24
21	16	v10,v10,v10	36	25	v21,v21,v21,v21,v21,v21	51	28	v23,v23,v23,v23,v23,v24,v24
22	19	v08,v09,v09,v09	37	22	v19,v19,v19,v20,v20	52	28	v23,v23,v23,v23,v24,v24,v24
23	16	v11,v12,v12	38	25	v20,v21,v21,v21,v21,v21	53	28	v23,v23,v23,v24,v24,v24,v24
24	19	v13,v13,v14,v14	39	22	v19,v20,v20,v20,v20	54	28	v23,v23,v24,v24,v24,v24,v24
25	22	v13,v13,v13,v13,v13	40	22	v20,v20,v20,v20,v20	55	28	v23,v24,v24,v24,v24,v24,v24
26	19	v15,v15,v15,v16	41	25	v19,v19,v19,v19,v19,v21			

表２の編成が理論的にどの程度の本選出場数となるのかの解析結果を表３に示す。表３では，それぞれの予選出場数 N に応じた本選出場チーム数の期待値，チーム数の上限と下限に対する理論値を示している。しかし，チーム数の下限の確率がきわめて小さい場合はそれを考慮することは現実的ではない。実際にいくつかの N に対する本選出場チーム数に関する確率分布は図３のような形状を持つので，表３では確率分布が95％である範囲（下限側2.5％と上限側2.5％の境界値），すなわち，$\arg_x P(N_2 < x)=0.025$，$\arg_x P(N_2 > x)=0.975$と本選出場チーム数 N_2 の目標値から，$P(N_2 < 8)$，$P(8 \leq N_2 \leq 12)$，$P(N_2 > 12)$の値を示している。この分布を図４にグラフ化した。

6　この条件は，予選出場チーム数をブロックに分割するための条件と，審査員数の制約としてブロック数の上限を７と設定することを意味する。

7　延べ数には本選審査員数 J_2 を含む。

表３：N ごとの N_2 に関する理論値

N	N_2 期待値	min	Max	$P(-8)$	$P(8-12)$	$P(12-)$	$P=2.5\%$	$P=97.5\%$
5	5	5	5	1	0	0	5	5
6	6	6	6	1	0	0	6	6
7	7	7	7	1	0	0	7	7
8	8	8	8	0	1	0	8	8
9	9	9	9	0	1	0	9	9
10	10	10	10	0	1	0	10	10
11	11	11	11	0	1	0	11	11
12	10.125	9	12	0	1	0	9	12
13	9.6689342	7	13	0.0008779	0.9981008	0.0010212	8	11
14	10.4489796	6	14	0.0003836	0.9860583	0.013558	9	12
15	10.8	6	15	0.0002835	0.9567054	0.043011	9	13
16	10.9375	8	14	0	0.9483857	0.0516142	9	13
17	10.9697959	7	14	0.0002233	0.9323514	0.0674253	9	13
18	9.75	6	15	0.0195142	0.9751354	0.0053506	8	12
19	9.4897959	5	14	0.0333168	0.9643405	0.0023427	7	12
20	10	5	15	0.0089925	0.9820151	0.0089925	8	12
21	10.4693878	4	17	0.0076357	0.9422579	0.0501063	8	13
22	11	7	15	0.000648	0.911704	0.087648	9	13
23	10.9406888	5	17	0.0038261	0.877456	0.1187177	8	14
24	9.5102041	4	14	0.0506694	0.9424088	0.0069217	7	12
25	10	5	15	0.0283809	0.9432383	0.0283809	7	13
26	9.2293878	3	16	0.1041309	0.8873167	0.0085525	7	12
27	9.5297959	4	16	0.0599633	0.9294981	0.0105386	7	12
28	9.9591837	3	18	0.0453717	0.9130698	0.0415586	7	13
29	10.2595918	4	17	0.0233813	0.9233487	0.05327	8	13
30	10.56	6	17	0.0100602	0.9215908	0.0683489	8	13
31	9.8625	3	17	0.0498413	0.9155653	0.0345935	7	13
32	10.125	4	16	0.0289147	0.9288974	0.0421879	7	13
33	10.53125	3	19	0.0223686	0.8756249	0.1020065	8	14
34	10.79375	4	18	0.0122095	0.8621686	0.125622	8	14
35	9.3877551	1	18	0.1175861	0.8575373	0.0248768	6	12
36	9.3333333	2	17	0.107035	0.8791957	0.0137693	6	12
37	9.8201531	1	18	0.0786087	0.87	0.0513913	7	13
38	9.8715278	2	18	0.0630835	0.8937135	0.043203	7	13
39	10.252551	2	19	0.0514737	0.8560524	0.092474	7	14
40	10.46875	2	19	0.04136	0.8397687	0.1188712	7	14
41	10.9433107	2	20	0.0212946	0.7982019	0.1805037	8	14
42	9.0612245	1	17	0.1704211	0.8125136	0.0170652	6	12
43	8.3523597	0	17	0.3078939	0.6851304	0.0069757	5	12
44	8.5006378	0	17	0.2785015	0.7123395	0.0091592	5	12
45	8.6489158	0	17	0.250857	0.7372446	0.0118983	5	12
46	8.7971939	0	17	0.225023	0.7596832	0.0152939	5	12
47	8.9454719	0	18	0.2010294	0.7795171	0.0194535	6	12
48	9.09375	1	18	0.1788773	0.796632	0.0244904	6	12
49	9.5714286	1	19	0.1256253	0.8209683	0.0534063	6	13
50	9.7197066	1	19	0.1104444	0.8258459	0.0637099	6	13
51	9.8679847	1	19	0.0967542	0.827845	0.0754006	6	13
52	10.0162628	1	19	0.0844674	0.8269844	0.0885482	6	14
53	10.1645408	1	20	0.0734895	0.8233045	0.1032057	7	14
54	10.3128189	1	20	0.063725	0.8168671	0.1194078	7	14
55	10.4610969	1	20	0.0550771	0.8077543	0.1371689	7	14

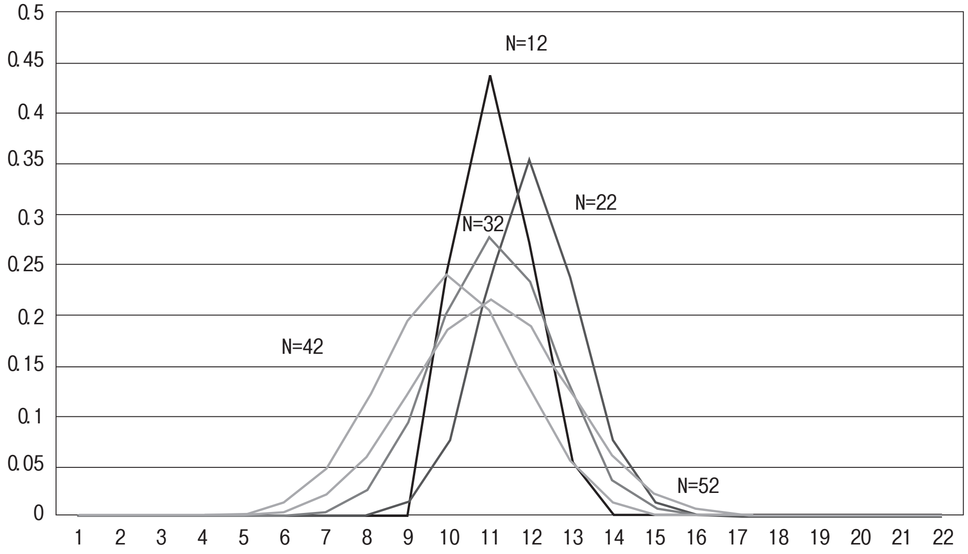


図3：N=12, 22, 32, 42, 52 における $P(N_2)$ の確率分布

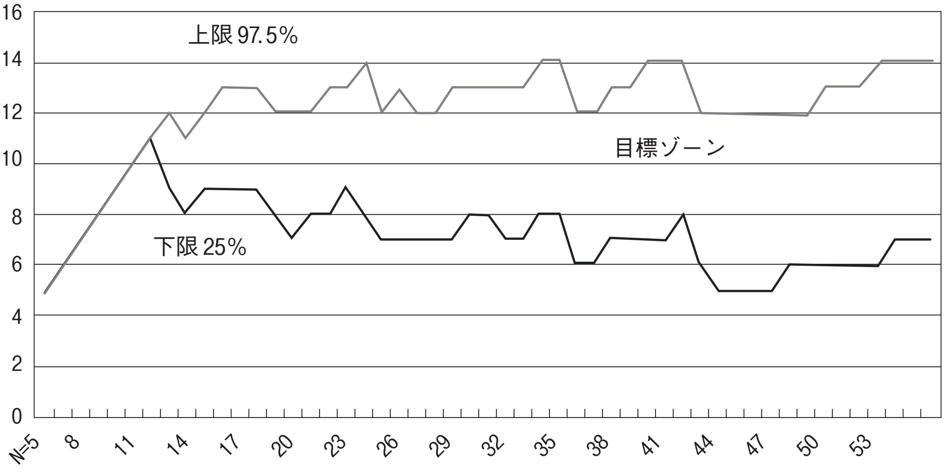


図4(a)：確率分布が95% である範囲

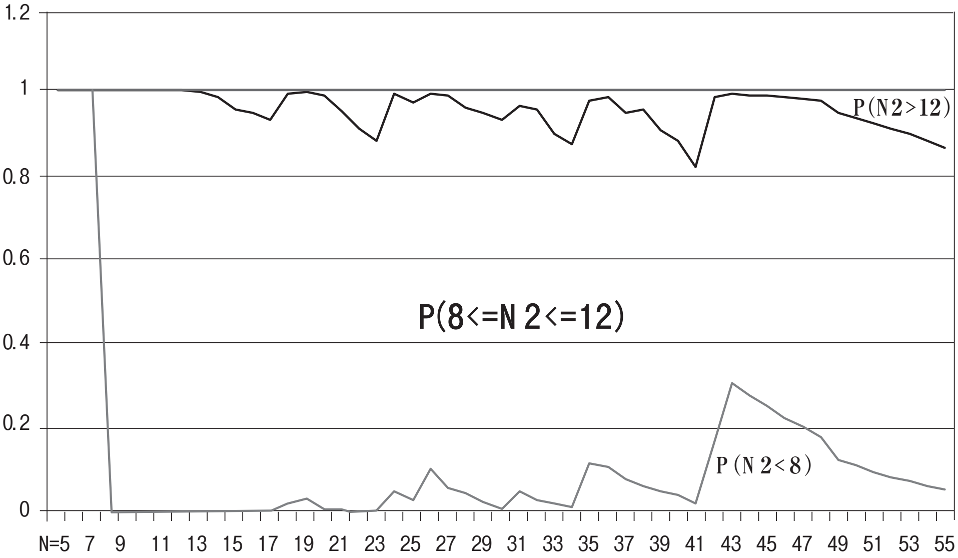


図４(b)：N_2 が目標値以内となる範囲

これらの結果から，予選審査の制度設計としておおむね妥当であることが分かる。しかし，本節の議論は，理論解析の必要からチームの実力が互いに等しいという仮定で行っている。次節ではそれを一般化したい。

５　シミュレーションによる評価

前節で予選のルールが設計された。本節ではこのルールの評価を行いたい。この問題で評価とは何を意味するのか整理すると，ブロック編成によって，実際に何チームが本選に出場したのかという正確性の側面と，ランダムに振り分けた実力の異なるチームの上位者がきちんと勝ち上がることができるのかという公平性の側面があげられるだろう。正確性と公平性の評価は，現実に近づけるためにチーム数に実力という概念を導入する[8]必要があり，これは解析を困難にする。そこでシミュレーションを行い評価することにする。

チーム i の実力g_i を $0 \leq g_i < 1$ とし，分布は一様分布を仮定する。また，審査員 j が認知するチーム i の実力を$g_i^j = g_i \pm \epsilon_i^j$とする。$\epsilon$ は認知誤差，あるいは，審査員の主観を表すホワイトノイズとし，i, j に独立に幅 h の一様分布とする。ここで，$h = 0.0$とは，全ての審査員がそれぞれのチームを同じに評価することを意味する。h が大きくなるにつれて，実力が伯仲するチームの評価が分かれることになる。本論では，$h = 0.0, 0.1, 0.2, 0.5$ についてシミュレートする。また，予選出場チーム数を $N = 20, 43, 52$とする。$N = 20$は期待値がちょうど10となるチーム数の少な

8　本稿で対象とするコンテストは評価が分かれるような定性性を有しているが，これをモデル化するのは困難であり，客観的な一般性を保持するために，本稿のような一元尺度で実力を定義することは妥当であると考えている。

い事例として，$N=43$は期待値が最も10から離れている事例として，$N=52$は期待値が10 に近くチーム数の多い事例として取り上げる。このときに，各ケースにおけるシミュレーションを異なる乱数種のもと，10,000回行い，そのときの正確性と公平性を次のように数値化して測定した。本選出場チーム数N_2の分布をグラフ化し，目標値$8 \leq N_2 \leq 12$を満足した試行割合を正確度とする。また，本選出場を果たせなかったチームのうち最も実力のあるチームの全体における順位kの分布をグラフ化し，各試行において$k/(N_2+1)$の平均を公平度とする[9]。

表４と図５にシミュレーション結果を示す。

表４：(N, h) 別の平均正確度と平均公平度

N	h	平均正確度	平均公平度
20	0	10	0.67925
20	0.1	10.0002	0.67264
20	0.2	9.9939	0.66562
20	0.5	9.9917	0.63482
43	0	6	0.55831
43	0.1	7.3933	0.59405
43	0.2	8.5256	0.57184
43	0.5	10.846	0.44266
52	0	7	0.51696
52	0.1	8.6603	0.55974
52	0.2	10.024	0.53328
52	0.5	12.7724	0.39736

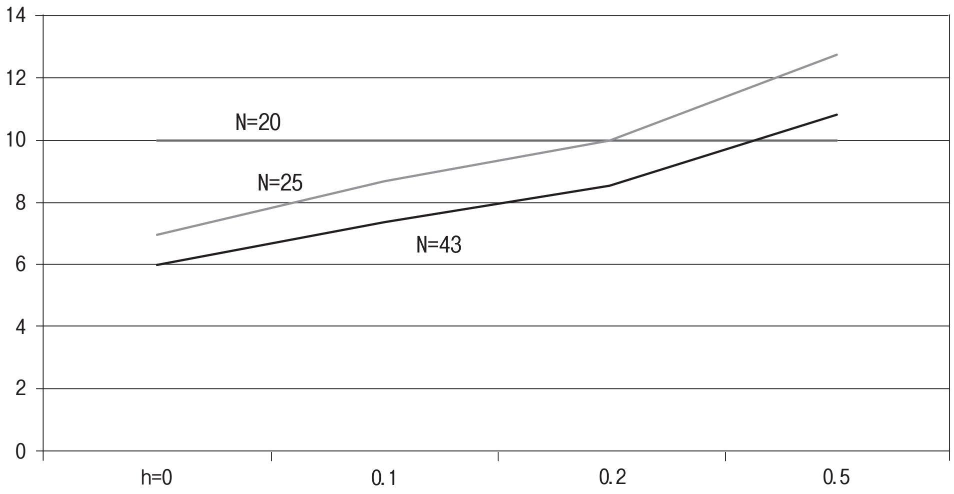
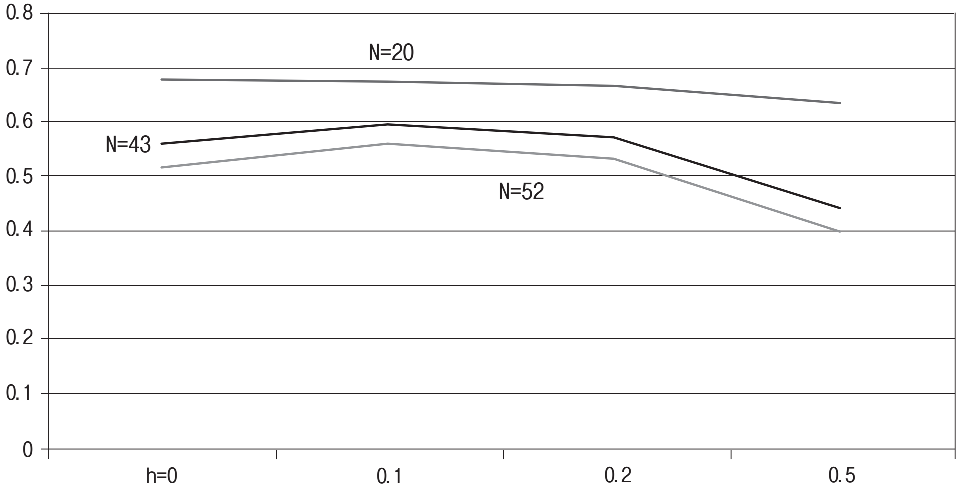
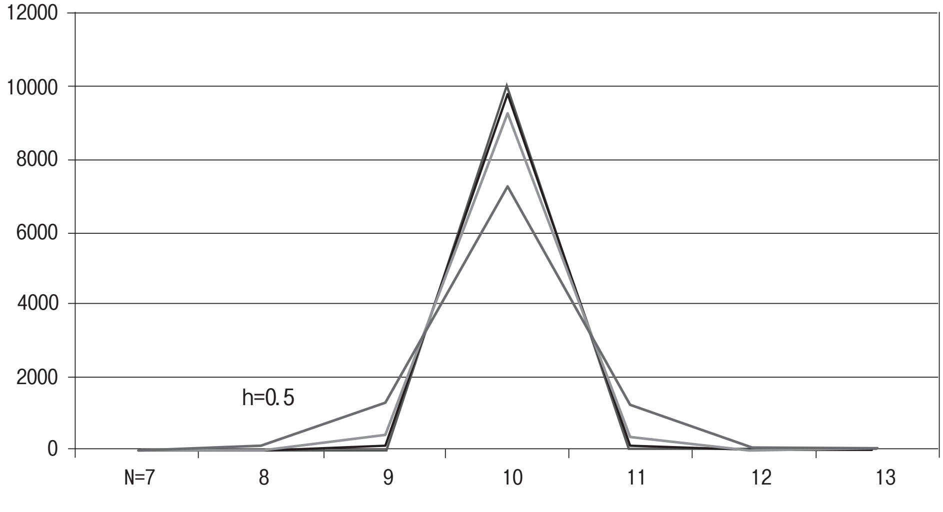


図５(a)：(N, h) 別の平均正確度

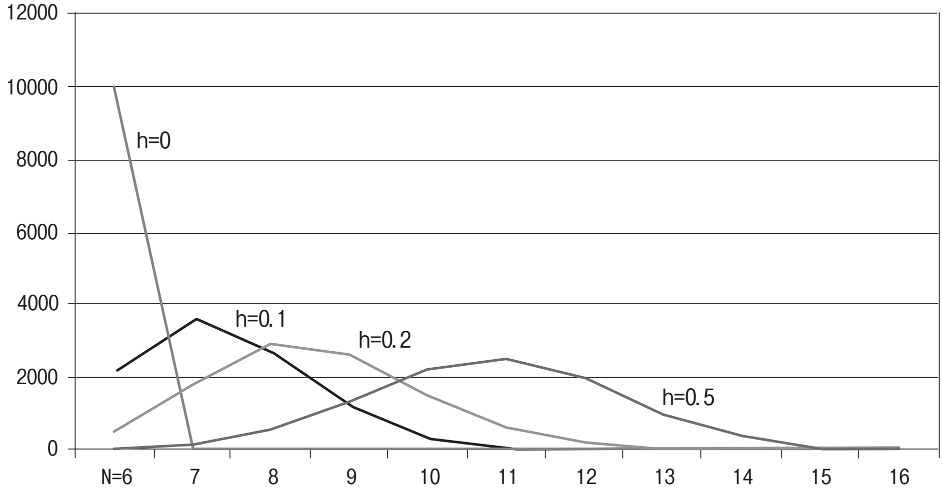
9　$1 \leq k \leq N_2+1$ であるから，公平度は $[1/N_2+1, 1]$ の範囲を取り公平であるほど値が高い。



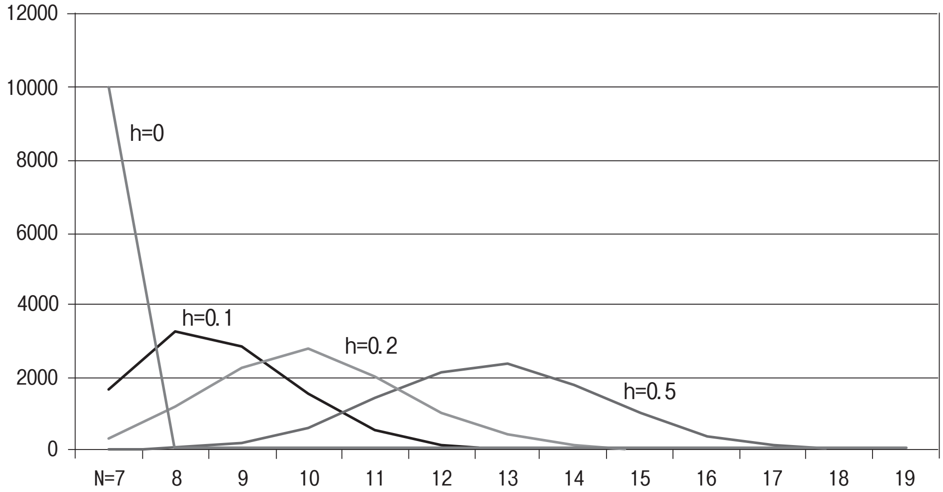
(b)：(*N*, ***h***) 別の平均公平度



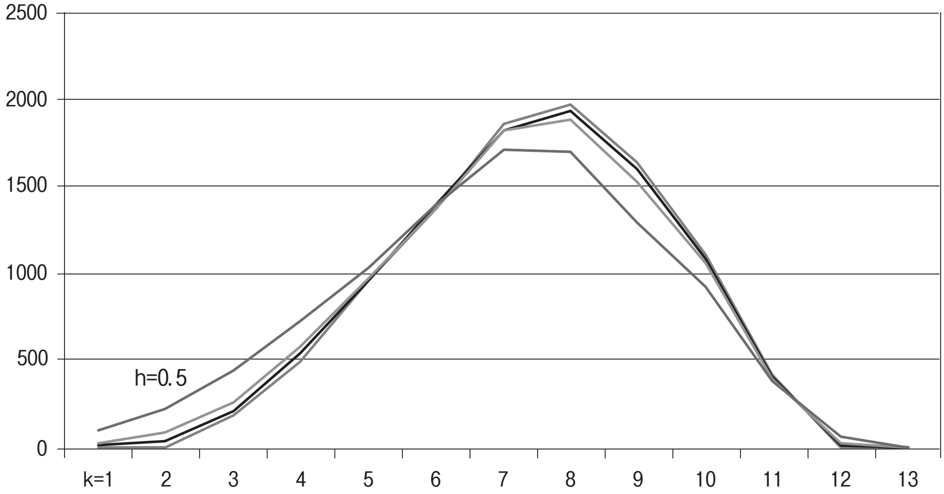
(c)：N＝20における ***h*** 別の N_2 の度数分布グラフ



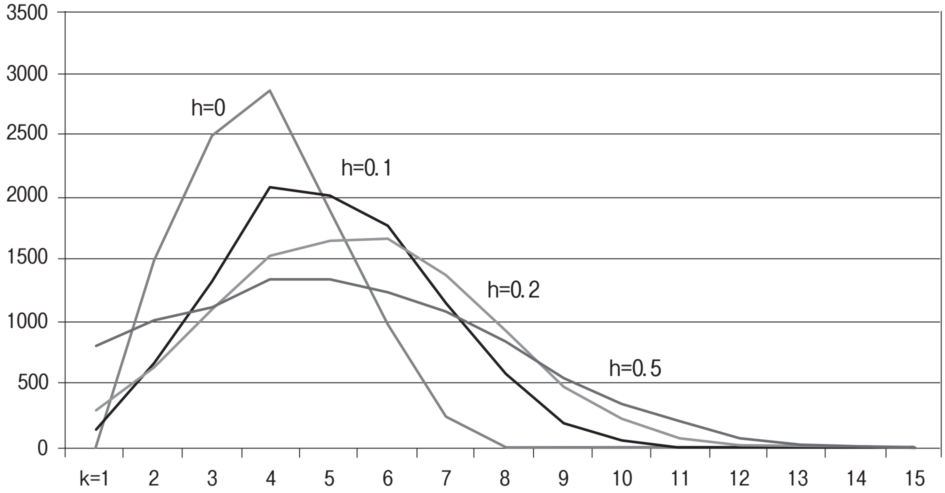
(d)：N＝43における h 別の N_2 の度数分布グラフ



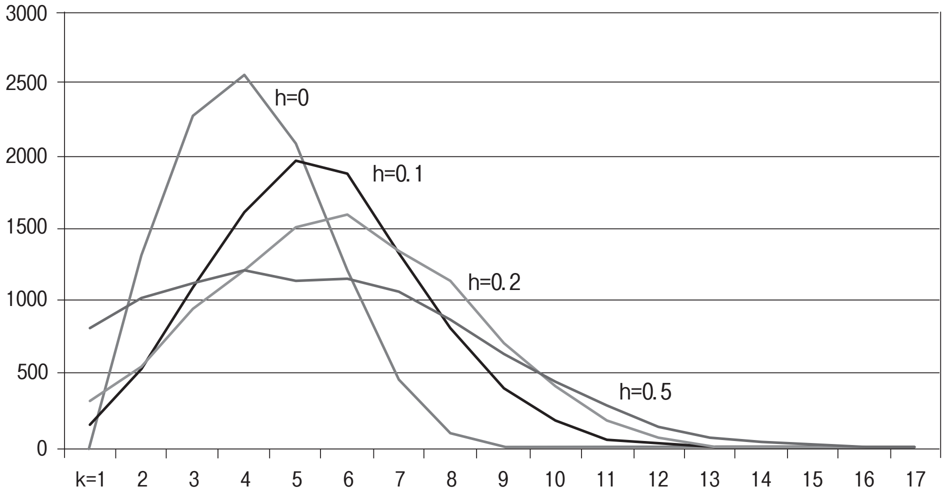
(e)：N＝52における h 別の N_2 の度数分布グラフ



(f)：N＝20における h 別の k の度数分布グラフ



(g)：N＝43における h 別の k の度数分布グラフ



(h)：N＝52における h 別の k の度数分布グラフ

まず，正確性について検討する。表4と図5から，$N=20$のときでは，審査が困難で結果が割れやすい（hが大きい）場合でも本選出場チーム数N_2は99.95% で目標値以内に収まるが，Nが大きくなるにつれて，審査の困難性がN_2を増大させていることが分かる[10]。特に$N=52$では，N_2の最頻値が目標値を上回るほどである。このことは，審査が困難であるほど，目標値を低めに設定するような制度設計が望ましいことを示唆している。しかし，審査困難性がある程度以内（$h=0.2$）ならば，目標値以内に収まる[11]ことから，審査結果のばらつきに関する先行データが制度の質を向上させる重要なファクターになることを意味している。

次に，公平性について検討する。まず$h=0.0$における状態を考える。これは審査が審査員の属性によらず一定でも生じる不公平度であるから，ブロック分けそのものに内在する不公平度といえる。それぞれ$N=$（20, 43, 52）のときは（0.68, 0.56, 0.52）となる。これはNに関わらずランダム状態における平均公平度が（0.18, 0.39, 0.47）程度である[12]ことからすると，Nの増加につれて公平度が低まることが分かる。次に，審査の困難性の影響を考える。$N=20$のときでは，審査が困難性に関わらず0.65程度の平均公平度となり，安定的に公平な状態を維持しているといえる。Nが大きくなっても，hがある程度（0.2）以下ならば平均公平度も変わらない。さらに，図5(g)，(h)によれば，hが高いほどkの最頻値が上昇している。これは，審査が困難であるほど上位実力者の本選出場可能性が高くなることを意味している。この知見は理論解析では発見しづらいもので，提案する審査制度の頑健性を強く示しているものといえる。

10　$h=0.5$のとき$N=43$では84.79% が，$N=52$では43.56% が目標値以内に収まっているに過ぎない。

11　$N=52$では92.05% が目標値以内に収まる。

12　ランダム状態における平均公平度は$N/(N_2(N_2+1))$で近似できる。

6　ま　と　め

本稿では審査員に高々３段階のブロック内の相対審査を求める予選方式を持つコンテストの設計問題について，特定のケースを想定して検討した。はじめに，抽象的な選択問題に関する理論解を抽出した。しかし，この理論モデルは実力が等しい対象に限定した分析となるため非現実的な制約を持っている。そのため，実力やその認知に関してばらつきを持たせた一般化データによりシミュレーションを行い，本稿で提案した制度の結果を分析した。制度の評価として正確性と公平性という２つの指標を定義し，シミュレーション結果を測定した。その結果，提案制度の頑健的な安定性を確認した。特に，審査が困難であるほど，本選出場チーム数の目標値を下げるべきであるといった知見や同じく審査困難性が上位実力者の出場可能性を増加させる頑健性を有しているといった知見が得られた。今後の拡張としては，対象とした問題の一般化や，実際の制度運用に関する制度の改善があげられるだろう。また，絶対基準と相対基準を組み合わせたような評価を行うことも可能であるから，この件における制度の拡張も試みることができる。